Weight Initializing Techniques

① Uniform Distribution

② Xavier Initialization

③ Kaiming He Initialization

Key Points

① Weights should be small

② Weights should not be same

③ Weights should have good variance

⟶ Uniform Distribution

$W_{ij} \approx$ Uniform Distribution

$$\left[\underbrace{\frac{-1}{\sqrt{\text{input}}}}_{\min(a)}, \underbrace{\frac{1}{\sqrt{\text{input}}}}_{\max(b)}\right]$$

if input-size = 3

$$\left[\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]$$

⟶ Xavier / Glorot ~~Distribution~~ Initialization

> Xavier Normal Initialization

$$w_{ij} \approx \overbrace{N(0,\sigma)}^{\text{normal distribution}}$$

$$\sigma = \sqrt{\frac{2}{(\text{input-size} + \text{output-size})}}$$

> Xavier Uniform

$w_{ij} \approx$ Uniform Distribution $\left[\frac{-\sqrt{6}}{\sqrt{\text{input} + \text{output}}}, \frac{\sqrt{6}}{\sqrt{\text{input} + \text{output}}}\right]$

⟶ Kaiming He Initialization

> He Normal

$$w_{ij} = N(0,\sigma)$$

$$\sigma = \sqrt{\frac{2}{\text{Input}}}$$

> He Uniform

$w_{ij} \approx$ Uniform Distribution $\left[-\sqrt{\frac{6}{\text{Input}}}, \sqrt{\frac{6}{\text{Input}}}\right]$